

On the Differentiation of Heat Semigroups and Poisson Integrals

ANTON THALMAIER

Abstract: We give a version of integration by parts on the level of local martingales; combined with the optional sampling theorem, this method allows us to obtain differentiation formulae for Poisson integrals in the same way as for heat semigroups involving boundary conditions. In particular, our results yield Bismut type representations for the logarithmic derivative of the Poisson kernel on regular domains in Riemannian manifolds corresponding to elliptic PDOs of Hörmander type. Such formulae provide a direct approach to gradient estimates for harmonic functions on Riemannian manifolds.

1. Introduction

Let M be an n -dimensional smooth manifold and, for some $m \in \mathbb{N}$, let

$$A: M \times \mathbb{R}^m \rightarrow TM, \quad (x, e) \mapsto A(x)e,$$

be a homomorphism of vector bundles over M . Thus, $A \in \Gamma(\mathbb{R}^m \otimes TM)$, i.e., the map $A(x): \mathbb{R}^m \rightarrow T_x M$ is linear for $x \in M$, and $A(\cdot)e \in \Gamma(TM)$ is a smooth vector field on M for $e \in \mathbb{R}^m$. Consider the Stratonovich stochastic differential equation

$$dX = A(X) * dB + A_0(X) dt \tag{1.1}$$

where $A_0 \in \Gamma(TM)$ is an additional vector field, and B an \mathbb{R}^m -valued Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in \mathbb{R}_+})$ satisfying the usual completeness conditions. There is a partial flow $X_t(\cdot), \zeta(\cdot)$ associated to (1.1) (see [12] for details) such that for each $x \in M$ the process $X_t(x)$, $0 \leq t < \zeta(x)$, is the maximal strong solution to (1.1) with starting point $X_0(x) = x$, defined up to the explosion time $\zeta(x)$; moreover, using the notation $X_t(x, \omega) = X_t(x)(\omega)$ and $\zeta(x, \omega) = \zeta(x)(\omega)$, if

$$M_t(\omega) = \{x \in M : t < \zeta(x, \omega)\}$$

then there exists a set $\Omega_0 \subset \Omega$ of full measure such that for all $\omega \in \Omega_0$:

1991 *Mathematics Subject Classification.* 58G32, 60H10, 60H30.

Key words and phrases. Diffusion, heat semigroup, integration by parts, heat kernel, Poisson kernel.

- (i) $M_t(\omega)$ is open in M for each $t \geq 0$, i.e. $\zeta(\cdot, \omega)$ is lower semicontinuous on M .
- (ii) $X_t(\cdot, \omega): M_t(\omega) \rightarrow M$ is a diffeomorphism onto an open subset of M .
- (iii) The map $s \mapsto X_s(\cdot, \omega)$ is continuous from $[0, t]$ into $C^\infty(M_t(\omega), M)$ with its C^∞ -topology, for each $t > 0$.

The solution processes $X = X(x)$ to (1.1) are diffusions on M with generator

$$L = A_0 + \frac{1}{2} \sum_{i=1}^m A_i^2$$

where $A_i = A(\cdot)e_i \in \Gamma(TM)$, $i = 1, \dots, m$. Throughout this paper we assume that the system (1.1) is *non-degenerate*, i.e., $A(x): \mathbb{R}^m \rightarrow T_x M$ is surjective for each x , or equivalently that L is elliptic. This non-degeneracy provides a Riemannian metric on M such that $A(x)A(x)^*: T_x M \rightarrow T_x M$ is the identity on $T_x M$ for $x \in M$. In other words, $A(x)^*: T_x M \rightarrow \mathbb{R}^m$ defines an isometric inclusion for each $x \in M$, i.e.,

$$\langle u, v \rangle_{T_x M} = \langle A(x)^* u, A(x)^* v \rangle_{\mathbb{R}^m} \quad \text{for all } u, v \in T_x M.$$

With respect to this Riemannian metric, $L = \frac{1}{2}\Delta_M + Z$ where Z is of first order, i.e. a vector field on M . Standard examples are the *gradient Brownian systems* when M is immersed into some Euclidean space \mathbb{R}^m , and $A(x): \mathbb{R}^m \rightarrow T_x M$ is the orthogonal projection; for $A_0 = 0$ this construction gives Brownian motion on M with respect to the induced metric, see [5].

For $x \in M$, let $T_x X_t: T_x M \rightarrow T_{X_t(x)} M$ be the differential of $X_t(\cdot)$ at x (well-defined for all $\omega \in \Omega$ such that $x \in M_t(\omega)$) and $V_t = V_t(v) = (T_x X_t)v$ the derivative process to $X_t(\cdot)$ at x in the direction $v \in T_x M$. It is well-known that V on TM solves the formally differentiated SDE (1.1), i.e.,

$$dV = (T_X A) V * dB + (T_X A_0) V dt, \quad V_0 = v, \quad (1.2)$$

with the same lifetime as $X(x)$, if $v \neq 0$. Using the metric and the corresponding Levi-Civita connection on M , equation (1.2) is most concisely written as a covariant equation along X

$$DV = (\nabla A) V * dB + (\nabla A_0) V dt \quad (1.3)$$

(see [5]); by definition, (1.3) means

$$d\tilde{V} = //_{0,t}^{-1}(\nabla A) //_{0,t} \tilde{V} * dB + //_{0,t}^{-1}(\nabla A_0) //_{0,t} \tilde{V} dt$$

for $\tilde{V}_t = //_{0,t}^{-1} V_t$ where $//_{0,t}: T_{X_0} M \rightarrow T_{X_t} M$ is parallel transport along the paths of X .

We first assume completeness in (1.1), i.e. $\zeta(x) = \infty$ a.s. for each $x \in M$. Note that this does not necessarily imply the existence of a sample continuous version of the flow $\mathbb{R}_+ \times M \rightarrow M$, $(t, x) \mapsto X_t(x)$. For $f \in bC^1(M)$ (bounded C^1 functions with bounded first derivative) let

$$(P_t f)(x) = \mathbb{E}[f \circ X_t(x)], \quad x \in M, \quad (1.4)$$

be the semigroup associated to (1.1), and

$$P_t^{(1)}(df)_x v = \mathbb{E}[(df)_{X_t(x)} (T_x X_t) v], \quad v \in T_x M, \quad (1.5)$$

its formal derivative whenever the right-hand side exists. More generally, for a (bounded) differential form $\alpha \in \Gamma(T^* M)$ let

$$P_t^{(1)}(\alpha) = \mathbb{E}[X_t^* \alpha], \quad (1.6)$$

provided the right-hand side of (1.6) is well-defined; here $X_t^* \alpha$ is the pullback of α under the (random) map $X_t: M \rightarrow M$.

Further, for $x \in M$ and $I = [0, t]$ or $I = \mathbb{R}_+$ let

$$\mathbb{H}(I, T_x M) = \{ \gamma: I \rightarrow T_x M \text{ absolutely continuous, } \|\dot{\gamma}\| \in L^2(I, ds) \}$$

be the Cameron-Martin space and $\mathbb{H}_0(I, T_x M) = \{ \gamma \in \mathbb{H}(I, T_x M) : \gamma(0) = 0 \}$.

The following version of an integration by parts formula is a slight variation of a formula obtained by Elworthy-Li [6] (see also [3]); we use it to exemplify our approach to derivative formulae.

THEOREM 1.1 (Integration by parts formula) *Assume (1.1) to be complete and non-degenerate. Let $f \in bC^1(M)$. Then*

$$\mathbb{E} \left[(df)_{X_t(x)} (T_x X_t) h_t \right] = \mathbb{E} \left[(f \circ X_t(x)) \int_0^t \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \right] \quad (1.7)$$

for each bounded adapted process h with sample paths in $\mathbb{H}_0([0, t], T_x M)$ such that $\mathbb{E} \left[\sup_{0 \leq s \leq t} |d(P_{t-s} f)_{X_s(x)} (T_x X_s) h_s| \right] < \infty$, and with the additional property that $\int_0^r \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle$, $0 \leq r \leq t$, is a martingale.

Proof Let h be an adapted bounded process with $h_*(\omega) \in \mathbb{H}([0, t], T_x M)$, almost all ω . It will be shown in Lemma 2.1 below that

$$\begin{aligned} N_r &= d(P_{t-r} f)_{X_r(x)} (T_x X_r) h_r \\ &\quad - (P_{t-r} f)(X_r(x)) \int_0^r \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \end{aligned} \quad (1.8)$$

provides a local martingale for $0 \leq r \leq t$. The additional assumptions assure that N is even a martingale; the claim follows upon taking expectation. \square

Remark 1.2 A canonical choice for h in equation (1.7) is $h_s = (s/t) v$, $v \in T_x M$, or more generally, $h_s = (s \wedge \varepsilon / \varepsilon) v$ with some constant $0 < \varepsilon \leq t$. Then, under the assumptions of Theorem 1.1,

$$\mathbb{E} \left[(df)_{X_t(x)} (T_x X_t) v \right] = \mathbb{E} \left[(f \circ X_t(x)) \frac{1}{\varepsilon} \int_0^\varepsilon \langle (T_x X_s) v, A(X_s(x)) dB_s \rangle \right]. \quad (1.9)$$

In general, if h in Theorem 1.1 has the property that $h_t = v$, then we get $P_t^{(1)}(df)_x v$ for the left-hand side in (1.7) while the right-hand side represents $d(P_t f)_x v$ as will be shown in Theorem 2.4 below. Thus, in this case, $d(P_t f)_x = P_t^{(1)}(df)_x$ is already a consequence of (1.7).

Note that differentiating (1.4) by taking derivatives under the expectation requires differentiability of f . However, due to the smoothing property of the semigroup, $P_t f$ is already differentiable even if f is only measurable — a fact which is explained by formula (1.7) where the right-hand side does not involve any derivatives of f .

In case system (1.1) is explosive, the minimal heat semigroup associated to (1.1) is given by

$$(P_t f)(x) = \mathbb{E} \left[(f \circ X_t(x)) 1_{\{t < \zeta(x)\}} \right] \quad (1.10)$$

where differentiation under the integral is no longer possible even for smooth f . An appropriate generalization of (1.5) is

$$P_t^{(1)}(df)_x v = \mathbb{E} \left[(df)_{X_t(x)} (T_x X_t) v 1_{\{t < \zeta(x)\}} \right], \quad v \in T_x M. \quad (1.11)$$

From a stochastic point of view, there seems to be no obvious reason why (1.11) should be the derivative of (1.10) in the direction v , i.e.,

$$dP_t f = P_t^{(1)}(df). \quad (1.12)$$

Of course, formula (1.12) cannot hold for $f \equiv 1$ unless the system (1.1) is complete (non-explosive).

Even more fundamental problems occur when dealing with boundaries where the process needs to be stopped when exiting a given domain. The situation is best illustrated in the case of the Dirichlet problem. Suppose that D is an open (relatively compact) domain in M with $\bar{D} \neq M$. Let

$$u(x) = \mathbb{E}[f \circ X_{\tau(x)}(x)] \quad (1.13)$$

where $\tau(x)$ denotes the first exit time of $X(x)$ from D . Then u is differentiable (and L -harmonic) on D whereas $x \mapsto X_{\tau(x)}(x)$ is not even continuous with probability one. The non-continuity follows from purely topological reasons, since there is no continuous retraction of D to the boundary ∂D . Thus, there is definitely no way of differentiating (1.13) by taking derivatives under the integral.

In this paper we shall extend integration by parts and derivative formulae in various directions to cover situations where finite lifetime or stopping times resulting from boundary conditions are involved. Specifically, we develop formulae for the differentiation of (1.10) and (1.13) not involving any derivatives of f . Analogously to Bismut type formulae for the logarithmic derivative of the heat kernel, we get similar formulae for the Poisson kernel.

Our methods are inspired by the notion of quasiderivatives in the sense of Krylov [11]. The following fact is elementary but crucial for our approach: If a local martingale depends on a parameter and is differentiable with respect to this parameter in probability uniformly on compact time intervals, then its derivative is also a local martingale.

2. A basic formula for the derivative of a heat semigroup

We start by explaining our basic strategy for proving integration by parts and derivative formulae; see also [7]. Let X be again the partial flow associated to the non-degenerate system (1.1). Suppose that, for some $x \in M$, a process of the form $Y(x): Y_r(x) = a(r, X_r(x))$ provides a local martingale. We assume that $Y(x)$ is defined on a stochastic interval $[\sigma, \tau[$ such that

$$\{(r, X_r(x, \omega)) : \sigma(\omega) \leq r < \tau(\omega)\} \subset I \times M_0, \quad \text{almost all } \omega,$$

where $a: I \times M_0 \rightarrow \mathbb{R}$ such that $a(t, \cdot) \in C^1(M_0)$ for $t \in I$ with jointly continuous derivative $(t, x) \mapsto da(t, \cdot)_x$; here $I \subset \mathbb{R}_+$ is an interval and $M_0 \subset M$ open.

Such a situation is typically given when a is some C^2 time-space harmonic function so that $\partial_r a + La = 0$. In this paper we have mainly two cases in mind, namely

- (i) $\tau \equiv t$, and $a(r, y) = (P_{t-r} f)(y)$, for some bounded measurable f on M ,
- (ii) $\tau \equiv \tau(x)$, where $\tau(x)$ is the first exit time of $X(x)$ from a bounded domain D , $a(r, y) = u(y)$ with $u \in C^2(D)$ and $Lu = 0$.

LEMMA 2.1 (Integration by parts on the local martingale level) *Let $a(r, X_r(x))$, $\sigma \leq r < \tau$ (with $\sigma < \tau$ predictable stopping times) be a local martingale for some function a having the above properties. Then*

$$N_r = (da(r, \cdot))_{X_r(x)} (T_x X_r) h_r - a(r, X_r(x)) \int_0^r \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \quad (2.1)$$

($\sigma \leq r < \tau$) is also a local martingale for any bounded adapted process h with sample paths $h_\bullet(\omega) \in \mathbb{H}(I, T_x M)$ for almost all ω .

Proof For $r \geq 0$, let $H_r^\varepsilon: M \rightarrow M$ be the pathwise defined solution to

$$\begin{cases} \frac{\partial}{\partial \varepsilon} H_r^\varepsilon(x) = A(H_r^\varepsilon(x)) A(x)^* h_r \\ H_r^0(x) = x. \end{cases} \quad (2.2)$$

Set $X_r^\varepsilon(x) = X_r(H_r^\varepsilon(x))$; then in particular $X_r^0(x) = X_r(x)$. The perturbed process X^ε satisfies

$$dX^\varepsilon = A(X^\varepsilon) * dB + A_0(X^\varepsilon) dr + (TX_r) dH_r^\varepsilon$$

with $dH_r^\varepsilon = (\frac{\partial}{\partial r} H_r^\varepsilon) dr = \dot{H}_r^\varepsilon dr$, see [12]. Hence

$$dX^\varepsilon(x) = A(X^\varepsilon(x)) * [dB + A(X^\varepsilon(x))^* (T_{H_r^\varepsilon(x)} X_r) dH_r^\varepsilon(x)] + A_0(X^\varepsilon(x)) dr.$$

We observe that this is an SDE of the same type as (1.1) but with the perturbed driving process $dB^\varepsilon(x) = dB + A(X^\varepsilon(x))^* (T_{H_r^\varepsilon(x)} X_r) dH_r^\varepsilon(x)$. Roughly speaking, the next step is to compensate this perturbation by changing the measure according to Girsanov-Maruyama. More precisely, set

$$M_r^\varepsilon = - \int_0^r \langle A(X_s^\varepsilon)^* (T_{H_s^\varepsilon} X_s) \dot{H}_s^\varepsilon, dB_s \rangle \quad (2.3)$$

and $G_r^\varepsilon = \exp(M_r^\varepsilon - \frac{1}{2}[M^\varepsilon]_r)$. Then, for any stopping time $\rho < \tau$ with the property that the exponential $(G_{r \wedge \rho}^\varepsilon(x))_{r \geq 0}$ is a martingale, $B^\varepsilon(x)|_{[0, \rho]}$ is a Brownian motion on $[0, \rho]$ with respect to the measure $G_\rho^\varepsilon(x) \cdot \mathbb{P}$. Hence, by pathwise uniqueness of solutions to (1.1), if $Y_r(x) = a(r, X_r(x))$ is a (local) martingale on $[\sigma, \tau[$ then also $Y_r^\varepsilon(x) := a(r, X_r^\varepsilon(x)) G_r^\varepsilon(x)$ is a local martingale on $[\sigma, \tau[$, both with respect to the measure \mathbb{P} . Consequently, also

$$\frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} Y_r^\varepsilon(x) = (da(r, \cdot))_{X_r(x)} (T_x X_r) h_r + a(r, X_r(x)) \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} G_r^\varepsilon(x),$$

for $\sigma \leq r < \tau$, is a local martingale. Using $\dot{H}_s^0 = 0$ and $\frac{\partial}{\partial s} \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} H_s^\varepsilon = \dot{h}_s$, we get

$$\begin{aligned} \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} G_r^\varepsilon(x) &= - \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} \int_0^r \langle (T_{H_s^\varepsilon(x)} X_s) \dot{H}_s^\varepsilon(x), A(X_s^\varepsilon(x)) dB_s \rangle \\ &= - \int_0^r \langle (T_x X_s) \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} \dot{H}_s^\varepsilon(x), A(X_s(x)) dB_s \rangle \\ &= - \int_0^r \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle. \end{aligned} \quad (2.4)$$

Thus $N_r = \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} Y_r^\varepsilon(x)$ where N_r is defined by (2.1). This shows that N is a local martingale on $[\sigma, \tau[$. \square

We shall exploit Lemma 2.1 for various choices of transformations a and processes h . An essential observation is that (before taking expectations in (2.1) with an appropriate h) there is still the possibility of applying the optional sampling theorem to the local martingale (2.1). This fact allows one to deal with stopping times in the derivative formulae which take into account given boundary conditions.

Remark 2.2 In the notation of Krylov [11] the local martingale property of (2.1) means that $\xi_r := (T_x X_r) h_r$ is a *quasiderivative* of $X_r(x)$ in the direction $h_0 = v$, and $\xi_r^0 := -\int_0^r \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle$ its *adjoint process*.

Now, let $f \in B(M)$ (bounded measurable functions on M), and suppose that the local martingale N_r , $0 \leq r \leq t$, as given by (1.8), is already a martingale; moreover suppose that $h_0 = v \in T_x M$ and $h_t = 0$. Then $\mathbb{E} N_0 = \mathbb{E} N_t$, in other words,

$$d(P_t f)_x v = -\mathbb{E} \left[(f \circ X_t(x)) \int_0^t \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \right].$$

For instance, choosing $h_r = (1 - r \wedge \varepsilon/\varepsilon) v$ (where $0 \leq \varepsilon \leq t$), we get

$$d(P_t f)_x v = \mathbb{E} \left[(f \circ X_t(x)) \frac{1}{\varepsilon} \int_0^\varepsilon \langle (T_x X_s) v, A(X_s(x)) dB_s \rangle \right]$$

provided (1.8) is actually a martingale for this choice of h . The latter question can be reduced to integrability conditions on the derivative process $T_x X_r$ (see [6]). Obviously also an appropriate choice for h may be helpful to make (1.8) a martingale. We follow this idea in the next theorem.

THEOREM 2.3 *Assume that (1.1) is complete and non-degenerate. Let $f \in B(M)$, $t > 0$. Then*

$$d(P_t f)_x v = -\mathbb{E} \left[(f \circ X_t(x)) \int_0^t \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \right] \quad (2.5)$$

holds for any bounded adapted process h with sample paths in $\mathbb{H}(\mathbb{R}_+, T_x M)$ such that $(\int_0^{\tau(x) \wedge t} \|\dot{h}_s\|^2 ds)^{1/2} \in L^{1+\varepsilon}$ for some $\varepsilon > 0$, and the property that $h_0 = v$, $h_s = 0$ for all $s \geq \tau(x) \wedge t$; here $\tau(x)$ is the first exit time of $X(x)$ from an (arbitrarily chosen) relatively compact neighbourhood D of x .

Proof 1) We first assume $f \in C^1(M)$. In this case $\|d(P_s f)_x\|$ is bounded for $(s, x) \in [0, t] \times D$. Now, let N^{τ_0} be the local martingale (1.8) stopped at $\tau_0 = \tau(x) \wedge t$, i.e. $N_r^{\tau_0} := N_{\tau_0 \wedge r}$, $r \geq 0$. It suffices to show that N^{τ_0} is already a martingale. Namely, then

$$d(P_t f)_x v = \mathbb{E} N_{\tau_0} = -\mathbb{E} \left[(P_{t-\tau_0} f)(X_{\tau_0}(x)) \int_0^{\tau_0} \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \right],$$

and (2.5) follows from the Markov property $(P_{t-\tau_0} f)(X_{\tau_0}(x)) = \mathbb{E}^{\mathcal{F}_{\tau_0}} [f \circ X_t(x)]$. To check the martingale property of N^{τ_0} , we first note that $\sup_{0 \leq s \leq \tau_0} \|T_x X_s\| \in L^p$ for any $1 \leq p < \infty$: for this we may assume that M is already compact, since otherwise M can be modified outside of D without changing $T_x X_s$ for $s \leq \tau_0$; on compact manifolds the above integrability of the derivative process is well-known, e.g. [14]. Using this integrability of the derivative process the stochastic integral in (1.8) can be estimated by means of Burkholder-Davis-Gundy and Hölder's inequality:

$$\begin{aligned} \mathbb{E} \left| \int_0^\sigma \langle A(X_s(x))^* (T_x X_s) \dot{h}_s, dB_s \rangle \right| &\leq c \mathbb{E} \left(\int_0^{\tau_0} \|(T_x X_s) \dot{h}_s\|^2 ds \right)^{1/2} \\ &\leq c \left[\mathbb{E} \left(\sup_{0 \leq s \leq \tau_0} \|T_x X_s\| \right)^{\frac{1+\varepsilon}{\varepsilon}} \right]^{\frac{\varepsilon}{1+\varepsilon}} \cdot \left[\mathbb{E} \left(\int_0^{\tau_0} \|\dot{h}_s\|^2 ds \right)^{\frac{1+\varepsilon}{2}} \right]^{\frac{1}{1+\varepsilon}} \leq \text{const} < \infty \end{aligned}$$

for any stopping time $\sigma \leq \tau_0$. This verifies that N^{τ_0} is indeed a martingale.

2) In case $f \in B(M)$ only, we use that $\|d(P_s f)_x\|$ is bounded for $(s, x) \in [\varepsilon, t] \times D$ if $\varepsilon > 0$. However, depending on ε , the process h may be modified such that $h_s^\varepsilon = h_s$ for $s \leq \tau(x) \wedge (t - \varepsilon)$ and $h_s^\varepsilon = 0$ for $s \geq \tau(x) \wedge (t - \varepsilon/2)$, and cutting off appropriately between. Then the arguments used in 1) carry over to give (2.5) with h replaced by h^ε . Finally, the claimed formula follows by $\varepsilon \rightarrow 0$. \square

Note that in Theorem 2.3 the condition $(\int_0^t \|\dot{h}_s\|^2 ds)^{1/2} \in L^{1+\varepsilon}$ guarantees that $\int_0^{\tau(x) \wedge r} \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle$, $0 \leq r \leq t$, is a martingale, i.e. it assures the uniform integrability of

$$\left\{ \int_0^\sigma \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle : 0 \leq \sigma \leq \tau(x) \wedge t, \sigma \text{ stopping time} \right\}.$$

The same strategy as above can be applied to get derivative formulae for the heat semigroup in cases when (1.1) is explosive. More precisely, we have the following result.

THEOREM 2.4 *Let $(P_t f)(x) = \mathbb{E}[(f \circ X_t(x)) 1_{\{t < \zeta(x)\}}]$ be the minimal semigroup associated to (1.1) acting on bounded measurable functions $f: M \rightarrow \mathbb{R}$. Then*

$$d(P_t f)_x v = -\mathbb{E}\left[(f \circ X_t(x)) 1_{\{t < \zeta(x)\}} \int_0^{\tau(x) \wedge t} \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle\right] \quad (2.6)$$

for any bounded adapted process h with sample paths in $\mathbb{H}(\mathbb{R}_+, T_x M)$ such that $(\int_0^{\tau(x) \wedge t} \|\dot{h}_s\|^2 ds)^{1/2} \in L^{1+\varepsilon}$ for some $\varepsilon > 0$, and the property that $h_0 = v$, $h_s = 0$ for all $s \geq \tau(x) \wedge t$; here $\tau(x)$ is again the first exit time of $X(x)$ from some relatively compact neighbourhood D of x .

Proof If $\tau_0 = \tau(x) \wedge t$, then $(P_{t-\tau_0} f)(X_{\tau_0}(x)) = \mathbb{E}^{\mathcal{F}_{\tau_0}}[(f \circ X_t(x)) 1_{\{t < \zeta(x)\}}]$ by the Markov property. The rest of the proof of Theorem 2.3 carries over verbatim to give (2.6). \square

From (2.6) a Bismut type formula can be derived for the transition kernel associated to (1.1).

COROLLARY 2.5 *Let $p(t, \cdot, \cdot): M \times M \rightarrow \mathbb{R}_+$, $t > 0$, be the (smooth) heat kernel (with respect to the Riemannian volume) associated to (1.1) such that*

$$(P_t f)(x) = \mathbb{E}[(f \circ X_t(x)) 1_{\{t < \zeta(x)\}}] = \int_M p(t, x, y) f(y) \text{vol}(dy)$$

for any $f \in B(M)$. Then

$$d(\log p(t, \cdot, y))_x v = -\mathbb{E}\left[\int_0^{\tau(x) \wedge t} \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \middle| X_t(x) = y\right] \quad (2.7)$$

with h and $\tau(x)$ as in Theorem 2.4.

Proof Let $f \in C(M)$ of compact support. By the smoothness of $p(t, \cdot, \cdot)$ for $t > 0$, we can differentiate under the integral to obtain

$$d(P_t f)_x v = \int dp(t, \cdot, y)_x v f(y) \text{vol}(dy).$$

On the other hand, (2.6) may be rewritten as

$$d(P_t f)_x v = - \int p(t, x, y) f(y) \mathbb{E} \left[\int_0^{\tau_0} \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \middle| X_t(x) = y \right] \text{vol}(dy)$$

with $\tau_0 = \tau(x) \wedge t$. Comparing the last two equations proves formula (2.7). \square

We conclude this section with some remarks on differentiation under the expectation, more precisely, for instance, on the question under which conditions

$$dP_t f = P_t^{(1)}(df) \quad (2.8)$$

holds for $f \in bC^1(M)$. As above, let $(P_t f)(\cdot) = \mathbb{E}[(f \circ X_t(\cdot)) 1_{\{t < \zeta(\cdot)\}}]$ be the minimal semigroup associated to our (possibly explosive) system (1.1), whereas $P_t^{(1)}(df)$ is given by (1.11). To make $P_t^{(1)}(df)$ well-defined we assume that

$$(df)_{X_t(x)} (T_x X_t) v 1_{\{t < \zeta(x)\}} \in L^1(\mathbb{P}).$$

Fixing a bounded adapted process h such that $h_\bullet(\omega) \in \mathbb{H}([0, t], T_x M)$ for almost all ω , we know that

$$N_r \equiv N_r^{(h)} = d(P_{t-r} f)_{X_r(x)} (T_x X_r) h_r - (P_{t-r} f)(X_r(x)) \int_0^r \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \quad (2.9)$$

defines a local martingale on the stochastic interval $[0, t \wedge \zeta(x)]$. Crucial for formula (2.8) to hold are basically two things: firstly, $N^{(h)}$ is required to be a uniformly integrable martingale for certain choices of h , and secondly, we need to know that $d(P_t f)_x \rightarrow 0$ sufficiently fast, as $x \rightarrow \infty$ in the one-point-compactification $M \cup \{\infty\}$ of M .

THEOREM 2.6 *Let h be a bounded adapted process with paths in $\mathbb{H}_0([0, t], T_x M)$ such that $h_t = v$. Given the above setting, suppose that a.s.*

$$d(P_{t-r} f)_{X_r(x)} (T_x X_r) v \rightarrow 0 \quad \text{on } \{\zeta(x) \leq t\} \text{ as } r \nearrow \zeta(x). \quad (2.10)$$

If $N^{(h)}$ defines a martingale, then

$$P_t^{(1)}(df)v = \mathbb{E} \left[(f \circ X_t(x)) 1_{\{t < \zeta(x)\}} \int_0^t \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \right].$$

If $N^{(h')}$ defines a martingale where $h'_s = v - h_s$, then

$$d(P_t f)_x v = \mathbb{E} \left[(f \circ X_t(x)) 1_{\{t < \zeta(x)\}} \int_0^t \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \right].$$

Proof The assertions follow from evaluating $\mathbb{E}[N_0] = \mathbb{E}[\lim_{r \nearrow t \wedge \zeta(x)} N_r]$, first for $N = N^{(h)}$, and then for $N = N^{(h')}$. \square

Remark 2.7 Keeping the notations of Theorem 2.6, we get the following criterion for $d(P_t f) = P_t^{(1)}(df)$. Suppose that assumption (2.10) holds (which is void for conservative systems). Moreover, suppose that

$$N_r \equiv N_r^{(v)} = d(P_{t-r} f)_{X_r(x)} (T_x X_r) v, \quad 0 \leq r \leq t \wedge \zeta(x),$$

is already a martingale. Note that $N^{(v)} \equiv N^{(h)}$ for $h \equiv v$. Then

$$d(P_t f)_x v = P_t^{(1)}(df)v,$$

which is seen again by taking expectations of $N^{(v)}$.

We want to stress that for explosive systems the domain of $P_t^{(1)}$ on 1-forms, i.e.

$$(P_t^{(1)}\alpha)_x v = \mathbb{E}[\alpha_{X_t(x)}(T_x X_t) v \mathbf{1}_{\{t < \zeta(x)\}}], \quad (2.11)$$

generally does not include bounded forms $\alpha \in \Gamma(T^*M)$, since $\|(T_x X_t)v\| \mathbf{1}_{\{t < \zeta(x)\}}$ will not be integrable in this case [13]. More precisely, we have the following.

Remark 2.8 Suppose that on a complete Riemannian manifold M relation (2.8) holds for all $f \in C^\infty(M)$ with compact support. If the derivative process has first moments, e.g., $\mathbb{E}\|(T_x X_t) \mathbf{1}_{\{t < \zeta(x)\}}\| < \infty$ for all x in some open set U and all $t \leq t_0$ where $t_0 > 0$, then the system (1.1) is already non-explosive [13]. Indeed, on any (geodesically) complete Riemannian manifold M one can construct an increasing sequence (f_n) of nonnegative smooth functions of compact support such that $f_n \nearrow 1$ and $\|df_n\|_\infty \leq 1/n$ for each n . Then $d(P_t f_n) \rightarrow d(P_t 1)$ on M , by standard Schauder type estimates. However, for $x \in U$ and $t \leq t_0$,

$$\|d(P_t f_n)_x\| = \|P_t^{(1)}(df_n)_x\| \leq \|df_n\|_\infty \mathbb{E}\|(T_x X_t) \mathbf{1}_{\{t < \zeta(x)\}}\| \rightarrow 0.$$

Thus, if $u(t, x) = (P_t 1)(x) = \mathbb{P}\{t < \zeta(x)\}$ then $u \equiv 1$ on $[0, t_0] \times U$, and finally $P_t 1 \equiv 1$.

3. The differentiation of Poisson integrals

We consider again a non-degenerate SDE of the type (1.1). Let $D \subset M$ be a nonvoid relatively compact open subset with $\bar{D} \neq M$, and

$$\tau(x) = \inf\{t \geq 0 : X_t(x) \notin D\}$$

the first exit time of X from D when started at $x \in D$. For $\varphi \in C(\partial D)$ let $u(x) = \mathbb{E}[\varphi \circ X_{\tau(x)}(x)]$. Then $Lu = 0$ on D .

THEOREM 3.1 *Assume that (1.1) is non-degenerate. Let $u(x) = \mathbb{E}[\varphi \circ X_{\tau(x)}(x)]$. Then*

$$(du)_x v = -\mathbb{E}\left[(\varphi \circ X_{\tau(x)}(x)) \int_0^{\tau(x)} \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle\right] \quad (3.1)$$

for any bounded adapted process h such that $h_\bullet \in \mathbb{H}(\mathbb{R}_+, T_x M)$, $h_0 = v$, and $h_s \equiv 0$ for $s \geq \tau(x)$, almost surely, provided $\int_0^{\tau(x) \wedge r} \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle$, $r \geq 0$, is a uniformly integrable martingale.

Proof Note that $u \circ X(x)$ is a martingale on $[0, \tau(x)]$, in particular a local martingale on $[0, \tau(x)[$. By Lemma 2.1, also

$$N_r = (du)_{X_r(x)}(T_x X_r) h_r - u \circ X_r(x) \int_0^r \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle$$

$(0 \leq r < \tau(x))$ is a local martingale. Since \bar{D} is compact, both u and du are bounded; moreover $\sup_{0 \leq s \leq \tau(x)} \|T_x X_s\| \in L^p$ for any $1 \leq p < \infty$. Using these properties it is easily checked that N is already a martingale on $[0, \tau(x)]$. The assertion follows then by taking expectations. \square

Note that, in the situation of Theorem 3.1, the process

$$\int_0^{\tau(x) \wedge t} \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle, \quad t \geq 0, \quad (3.2)$$

is a uniformly martingale if for instance $(\int_0^{\tau(x)} \|\dot{h}_s\|^2 ds)^{1/2} \in L^{1+\varepsilon}$ for some $\varepsilon > 0$. Using $\sup_{0 \leq s \leq \tau(x)} \|T_x X_s\| \in L^p$ for any $1 \leq p < \infty$, the stochastic integrals in (3.2)

can be estimated by means of Burkholder-Davis-Gundy and Hölder's inequality, as in the proof of Theorem 2.3.

COROLLARY 3.2 *Assume that $D \subset M$ is a nonvoid relatively compact open domain with $\bar{D} \neq M$ and with smooth boundary. Let $p: D \times \partial D \rightarrow \mathbb{R}_+$ be the (smooth) Poisson kernel (with respect to the induced surface measure σ on ∂D) so that*

$$\mathbb{P}\{X_{\tau(x)}(x) \in dz\} = p(x, z) \sigma(dz).$$

Then

$$d(\log p(\cdot, z))_x v = -\mathbb{E}\left[\int_0^{\tau(x)} \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \Big| X_{\tau(x)}(x) = z\right] \quad (3.3)$$

where h with sample paths in the Cameron-Martin space is as in Theorem 3.1.

Proof For $\varphi \in C(\partial M)$, let $u(x) = \mathbb{E}[\varphi \circ X_{\tau(x)}(x)]$. We differentiate u under the integral to obtain

$$(du)_x v = \int_{\partial M} dp(\cdot, z)_x v \varphi(z) \sigma(dz).$$

On the other hand, by rewriting (3.1) we get

$$(du)_x v = -\int p(x, z) \varphi(z) \mathbb{E}\left[\int_0^{\tau(x)} \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \Big| X_{\tau(x)}(x) = z\right] \sigma(dz).$$

Comparing the last two equations gives formula (3.3). \square

4. Remarks on the choice of the process h

Let D be an open (relatively compact) domain in M . Given $x \in D$ and $v \in T_x M$, most of our formulae require the choice of a bounded adapted process h with sample paths in $\mathbb{H}(\mathbb{R}_+, T_x M)$ such that, e.g., $h_0 = 0$ and $h_s = v$ for $s \geq \tau(x)$, and the property that $(\int_0^{\tau(x)} \|\dot{h}_s\|^2 ds)^{1/2} \in L^{1+\varepsilon}$ for some $\varepsilon > 0$; here $\tau(x)$ is the first exit time of X from D when starting at x . We describe a method of constructing such processes; see [15] for more details.

Suppose that D has smooth boundary. We take $f \in C^2(\bar{D})$ with $f|_{\partial D} = 0$ and $f > 0$ in D . Further fix $x \in D$ and write τ instead of $\tau(x)$. Consider the increasing process

$$T(t) = \int_0^t f^{-2}(X_s(x)) ds, \quad t \leq \tau,$$

and

$$\sigma(t) = \inf\{s \geq 0 : T(s) > t\}, \quad t \leq T(\tau).$$

Obviously $T(\sigma(t)) = t$ for $t \leq T(\tau)$, and $\sigma(T(t)) = t$ for $t \leq \tau$. Since $X_\bullet(x)$ is an L -diffusion with generator $L = \frac{1}{2}\Delta + Z$, the time-changed process $\tilde{X}_t(x) = X_{\sigma(t)}(x)$ is an \tilde{L} -diffusion where $\tilde{L} = f^2 L$. The following lemma shows that \tilde{L} -diffusions on D have infinite lifetime. As a consequence, we get $T(\tau) = \infty$ a.s.

LEMMA 4.1 Let \tilde{X} be an \tilde{L} -diffusion on D with $\tilde{X}_0 = x \in D$. Then

$$\tilde{\tau} \equiv \inf\{s \geq 0 : \tilde{X}_s \in \partial D\} = \infty, \text{ a.s.}$$

Proof Recall that \tilde{X} is a \tilde{L} -diffusion if, for any $\varphi \in C^2(D)$,

$$\varphi \circ \tilde{X}_t - \varphi(x) - \int_0^t \tilde{L}\varphi(\tilde{X}_s) ds, \quad 0 \leq t \leq \tilde{\tau},$$

is local martingale. For $n \geq 1$, let $\tau_n = \inf\{s \geq 0 : f(\tilde{X}_s) \leq 1/n\}$, and choose $n_0 \geq 1$ such that $f(x) \geq 1/n_0$. Note that $\tilde{L}f^{-1} = -Lf + f^{-1}\|\text{grad } f\|^2 \leq cf^{-1}$ for some constant $c = c(f)$. Thus

$$E^x f^{-1}(\tilde{X}_{t \wedge \tau_n}) \leq f^{-1}(x) e^{ct}, \quad t \geq 0, \quad n \geq n_0.$$

But $\mathbb{E}[f^{-1}(\tilde{X}_{t \wedge \tau_n})] \geq n \mathbb{P}\{\tau_n < t\}$, hence

$$\mathbb{P}\{\tau_n < t\} \leq n^{-1} f^{-1}(x) e^{ct}.$$

Therefore, $\mathbb{P}\{\tilde{\tau} < t\} = 0$ for any $t \geq 0$. This proves the Lemma. \square

Now, for fixed $t_0 > 0$, let

$$h_s = v \frac{1}{t_0} \int_0^s f^{-2}(X_r(x)) 1_{\{r < \sigma(t_0)\}} dr. \quad (4.1)$$

Then, for $s \geq \sigma(t_0)$,

$$h_s = h(\sigma(t_0)) = v \frac{1}{t_0} \int_0^{\sigma(t_0)} f^{-2}(X_r(x)) dr = v.$$

It remains to verify that $(\int_0^{\sigma(t_0)} \|\dot{h}_s\|^2 ds)^{1/2} \in L^{1+\varepsilon}$ for some $\varepsilon > 0$. For instance, we may take $\varepsilon = 1$. Obviously,

$$\int_0^{\sigma(t_0)} \|\dot{h}_s\|^2 ds = \|v\|^2 \frac{1}{t_0^2} \int_0^{\sigma(t_0)} f^{-4}(X_s(x)) ds = \|v\|^2 \frac{1}{t_0^2} \int_0^{t_0} f^{-2}(X_{\sigma(s)}(x)) ds.$$

Recall that $X_{\sigma(s)}(x) = \tilde{X}_s(x)$ and $d f^{-2}(\tilde{X}_s) = dN_s + \tilde{L}f^{-2}(\tilde{X}_s) ds$ where (N_s) is a local martingale. But $\tilde{L}f^{-2} = f^2 L f^{-2} = f^2 [3f^{-4}\|\text{grad } f\|^2 - 2f^{-3}Lf] \leq cf^{-2}$ for some constant $c = c(f)$, hence

$$\mathbb{E}[f^{-2}(\tilde{X}_s(x))] \leq f^{-2}(x) e^{cs},$$

and thus, as claimed,

$$\mathbb{E}\left[\int_0^{\sigma(t_0)} f^{-4}(X_s(x)) ds\right] = \mathbb{E}\left[\int_0^{t_0} f^{-2}(\tilde{X}_s(x)) ds\right] \leq f^{-2}(x) \frac{e^{ct_0} - 1}{t_0} < \infty.$$

Note that the process h , as defined in (4.1), depends on t_0 and f . For any $t_0 > 0$ and any $f \in C^2(\bar{D})$ with $f|_{\partial D} = 0$ and $f > 0$ in D , formula (4.1) gives a process with the required properties.

5. Extensions to closed differential forms

For the sake of simplicity we restrict ourselves to the case when the system (1.1) defines Brownian motion on (M, g) ; generalizations to h -Brownian motion (see [6])

for instance are straightforward. Let (M, g) be a complete Riemannian manifold and

$$\dots \xrightarrow{d} \Gamma(\Lambda^p T^* M) \xrightarrow{d} \Gamma(\Lambda^{p+1} T^* M) \xrightarrow{d} \dots$$

its deRham complex. Denote by Δ the deRham-Hodge-Laplace operator defined as the L^2 -closure of the operator $-(d + \delta)^2$ on compactly supported elements of $\Gamma(\Lambda^\bullet T^* M) = \bigoplus_{p \geq 0} \Gamma(\Lambda^p T^* M)$. Let $\text{dom } \Delta \subset L^2\text{-}\Gamma(\Lambda^\bullet T^* M)$ be the domain of Δ , and $\Delta_1 = \Delta|_{\text{dom } \Delta \cap L^2\text{-}\Gamma(T^* M)}$ the restriction of Δ to 1-forms; note that $L^2\text{-}\Gamma(\Lambda^\bullet T^* M)$ are the L^2 -sections of $\Lambda^\bullet T^* M$, in contrast to $\Gamma(\Lambda^\bullet T^* M)$ which denotes the smooth sections. By the spectral theorem, there is a smooth semigroup $P_t = e^{(1/2)t\Delta_1}$ on $L^2\text{-}\Gamma(T^* M)$ solving the heat equation

$$\frac{\partial}{\partial t} P_t \alpha = \frac{1}{2} \Delta P_t \alpha. \quad (5.1)$$

Note that $P_t \alpha \in \Gamma(T^* M)$ for $\alpha \in L^2\text{-}\Gamma(T^* M)$ due to elliptic regularity. For a differential form $\alpha \in \Gamma(T^* M)$, let $\int_X \alpha$ be the Stratonovich integral, and $\int_X^{(1)} \alpha$ the Itô integral of α along $X = X(x)$ (see [10]). Recall that

$$\int_X \alpha = \int_X^{(1)} \alpha + \frac{1}{2} \int \nabla \alpha (dX, dX) = \int_X^{(1)} \alpha - \frac{1}{2} \int \delta \alpha (X_s(x)) ds.$$

In our situation, we have $(\int_X \alpha)_r = \int_0^r \alpha_X(*dX) = \int_0^r \alpha_{X_s(x)}(A(X_s(x)) * dB_s)$, and $(\int_X^{(1)} \alpha)_r = \int_0^r \alpha_{X_s(x)}(A(X_s(x)) dB_s)$. Analogously, for the “time-dependent” differential forms $P_{t-} \alpha$, we set $(\int_X P_{t-} \alpha)_r = \int_0^r (P_{t-s} \alpha)_{X_s(x)}(A(X_s(x)) * dB_s)$, and $(\int_X^{(1)} P_{t-} \alpha)_r = \int_0^r (P_{t-s} \alpha)_{X_s(x)}(A(X_s(x)) dB_s)$. The following theorem is along the lines of Elworthy-Li [6].

THEOREM 5.1 *Suppose that the system (1.1) defines Brownian motion on (M, g) , possibly with finite lifetime. Let $\alpha \in L^2\text{-}\Gamma(T^* M) \cap \text{dom } \Delta$ with $d\alpha = 0$. Then*

$$N_r = (P_{t-r} \alpha)_{X_r(x)} T_x X_r h_r - \left(\int_X^{(1)} P_{t-} \alpha \right)_r \int_0^r \langle T_x X_s \dot{h}_s, A(X_s(x)) dB_s \rangle, \quad (5.2)$$

$r \in [0, t] \cap [0, \zeta(x)[$, is a local martingale for any adapted bounded process h such that $h_\bullet(\omega) \in \mathbb{H}([0, t], T_x M)$, almost all ω .

Proof Again the situation is reduced to Lemma 2.1. By a standard localization argument, e.g., [10], Lemma (3.5), it is enough to check the local martingale property of N on stochastic intervals contained in sets of the form $\{X \in V_i\}$ where $(V_i)_{i \in I}$ forms an open covering of M . First, since α is closed, we get $dP_r \alpha = P_r d\alpha = 0$. Hence, for each $r_0 \in [0, t]$ and $x_0 \in M$ there is an open neighbourhood V of x_0 such that $P_r \alpha = da_r$ on V for all r in some open interval I about $t - r_0$; moreover a_r can be chosen such that $(\frac{\partial}{\partial r} + \frac{1}{2} \Delta^M) a_r = 0$ on $I \times V$. We may assume that $(r, x) \mapsto a_r(x)$ is bounded on $I \times V$. Now, let $[\sigma, \tau[\subset I$ be a stochastic interval such that $X|_{[\sigma, \tau[}$ takes values in V , then on $[\sigma, \tau[$

$$\begin{aligned} d\left(\int_X^{(1)} P_{t-} \alpha\right) &= (da_{t-r})_{X_r(x)} * dX_r + \frac{1}{2} (\Delta a_{t-r})(X_r(x)) dr \\ &= (da_{t-r})_{X_r(x)} * dX_r + (\partial_r a_{t-r})(X_r(x)) dr = d(a_{t-r} \circ X_r(x)). \end{aligned}$$

Thus, $N|_{[\sigma, \tau[}$ is a local martingale by Lemma 2.1. \square

COROLLARY 5.2 *For any harmonic 1-form $\alpha \in \Gamma(T^* M) \cap L^2$ the process*

$$N_r = \alpha_{X_r(x)} (T_x X_r h_r) - \left(\int_X \alpha \right)_r \int_0^r \langle T_x X_s \dot{h}_s, A(X_s(x)) dB_s \rangle \quad (5.3)$$

defines a local martingale, $0 \leq r < \zeta(x)$.

Proof Obviously $\int_X \alpha = \int_X^{(1)} \alpha$ for harmonic 1-forms α . \square

Remark 5.3 Let $Q_r \alpha = -\frac{1}{2} \int_0^r P_s(\delta \alpha) ds$. Then, if $\zeta(x) = \infty$ a.s., we get

$$\left(\int_X^{(I)} P_{t-} \cdot \alpha \right)_t = \left(\int_X \alpha \right)_t - (Q_t \alpha)(x), \quad (5.4)$$

as can be seen by applying Itô's formula to $(Q_{t-s} \alpha)(X_s(x))$, $0 \leq s \leq t$. With the help of identity (5.4) it is straightforward to recover the corresponding formulae for 1-forms in [7] from Theorem 5.1.

6. Some variations of the differentiation formulae

In this section we rewrite our basic differentiation formulae (2.6) and (3.1) in terms of the conditional derivative process, as defined by Elworthy-Yor [9]. The resulting formulae will be intrinsic in the sense that, for fixed x , the right-hand sides are given entirely in terms of the diffusion $X(x)$, starting at x ; they involve no longer the derivative flow which depends on the SDE (1.1), used to obtain the diffusion $X(x)$.

The idea is to filter out extraneous noise of the local martingale (2.1) by conditioning with respect to the smaller filtration generated by $X(x)$. More precisely, for $x \in M$, let

$$\mathcal{F}_r(x) := \mathcal{F}_r^{X(x)} \equiv \sigma\{X_s(x) : 0 \leq s \leq r\}. \quad (6.1)$$

For some given $v \in T_x M$ consider again the derivative process $V_r(v) = (T_x X_r) v$. Fix an $\mathcal{F}_\bullet(x)$ -stopping time τ such that $V(v)$ is integrable on $[0, \tau]$, i.e.,

$$\|V_r(v) 1_{\{r \leq \tau\}}\| \in L^1(\mathbb{P})$$

for each $r \geq 0$, and define a TM -valued process $W(v)$ along $X(x)$ by

$$W_r(v) := \mathbb{E}^{\mathcal{F}_r(x)}[(T_x X_r) v 1_{\{r \leq \tau\}}] \equiv //_{0,r} \mathbb{E}^{\mathcal{F}_r(x)} //_{0,r}^{-1} [(T_x X_r) v 1_{\{r \leq \tau\}}] \quad (6.2)$$

where $//_{0,r}: T_x M \rightarrow T_{X_r(x)} M$ denotes parallel transport along $X(x)$. Note that, instead of conditioning with respect to $\mathcal{F}_r(x)$ in (6.2), we may equivalently take expectations with respect to $\mathcal{F}_\tau(x)$, or $\mathcal{F}_\infty(x)$. Recall that $L = \frac{1}{2} \Delta_M + Z$ where $Z \in \Gamma(TM)$. Then, as in [9], it can be shown that $W(v)$ satisfies the covariant equation

$$\begin{cases} \frac{D}{dr} W_r(v) = -\frac{1}{2} \text{Ric}(W_r(v), \cdot)^\# + \nabla Z(W_r(v)) \\ W_0(v) = v \end{cases} \quad (6.3)$$

along $X(x)$ for $r \leq \tau$. (Without loss of generality we may assume that the Levi-Civita on M coincides with the Le Jan-Watanabe connection associated to (1.1), see [8]). Note that $\frac{D}{dr} W_r(v) = //_{0,r} \frac{d}{dr} //_{0,r}^{-1} W_r(v)$ by definition; moreover, if $w \in T_y M$, then $\text{Ric}(w, \cdot)^\# \in T_y M$ is determined by $\langle \text{Ric}(w, \cdot)^\#, z \rangle = \text{Ric}(w, z)$ for all $z \in T_y M$.

Let U be a horizontal lift of $X(x)$ to the orthonormal frame bundle $\pi: O(M) \rightarrow M$, and $Z = U_0 \int_U \vartheta$ the anti-development of $X(x)$ in $T_x M$ with respect to the Levi-Civita connection, see [10]; here $\vartheta \in \Gamma(T^* O(M) \otimes \mathbb{R}^n)$, $\vartheta_u = u^{-1} d\pi_u$, $u \in O(M)$,

is the canonical 1-form of the connection. Thus $Z_r = \int_0^r //_{0,s}^{-1} * dX_s(x)$. Let

$$\tilde{B}_r = \int_0^r //_{0,s}^{-1} A(X_s(x)) dB_s \quad (6.4)$$

be the martingale part of Z ; then $A(X_s(x))dB_s = //_{0,s} d\tilde{B}_s$. On the other hand, it is easily seen that \tilde{B} is a Brownian motion on $T_x M$, stopped at the lifetime $\zeta(x)$ of $X(x)$. The point is that by construction \tilde{B} is adapted to the filtration $\mathcal{F}_\bullet(x)$ generated by $X(x)$.

We return to the general situation of Lemma 2.1 and consider the local martingale

$$N_r = (da(r, \cdot))_{X_r(x)} (T_x X_r) h_r - a(r, X_r(x)) \int_0^r \langle (T_x X_s) \dot{h}_s, A(X_s(x)) dB_s \rangle \quad (6.5)$$

on a stochastic interval $[\sigma, \tau[$. Here $a: I \times M_0 \rightarrow \mathbb{R}$ (with $I \subset \mathbb{R}_+$ an interval and $M_0 \subset M$ open) is again a transformation such that the process $(r, X_r(x))$ on $[\sigma, \tau[$ takes its values almost surely in $I \times M_0$, and $a(\cdot, X_\bullet(x))$ defines a local martingale; for the required technical properties of a see the beginning of section 2.

LEMMA 6.1 *Let $a(r, X_r(x))$, $\sigma \leq r < \tau$ (with $\sigma < \tau$ predictable stopping times) be a local martingale for some function a as above. Suppose that h is a bounded process with sample paths $h_\bullet(\omega) \in \mathbb{H}(I, T_x M)$, almost all ω , which is already adapted to $\mathcal{F}_\bullet(x)$. Then*

$$\tilde{N}_r = (da(r, \cdot))_{X_r(x)} W_r(h_r) - a(r, X_r(x)) \int_0^r \langle W_s(\dot{h}_s), //_{0,s} d\tilde{B}_s \rangle \quad (6.6)$$

is a local martingale on $[\sigma, \tau[$; here $W(\cdot)$ is defined by (6.3) and the Brownian motion \tilde{B} is given by (6.4).

Proof By Lemma 2.1, the process N , as defined in (6.5), is a local martingale. Conditioning of N with respect to $\mathcal{F}_\bullet(x)$ gives the claim. \square

With the help of Lemma 6.1, i.e., by working with \tilde{N} instead of N , we can rewrite our basic formulae in an obvious way. For instance, given the assumptions of Theorem 2.4, formula (2.6) reads as

$$d(P_t f)_x v = -\mathbb{E} \left[(f \circ X_t(x)) 1_{\{t < \zeta(x)\}} \int_0^{t \wedge \tau(x)} \langle W_s(\dot{h}_s), //_{0,s} d\tilde{B}_s \rangle \right]. \quad (6.7)$$

Analogous considerations apply to formula (3.1).

Brownian motions on manifolds are often constructed via stochastic development of Euclidean Brownian motion, see [5], [10]. We conclude this section with some comments how our results are easily adapted to this situation.

For a Riemannian manifold (M, g) denote by L_1, \dots, L_m the canonical horizontal vector fields on the orthonormal base bundle $\pi: O(M) \rightarrow M$ over M , given by $L_i(u) = \mathbf{h}_u(u e_i)$ where $\mathbf{h}: \pi^* TM \rightarrow TO(M)$ is the horizontal lift induced by the Levi-Civita connection on M . Let $H \subset TO(M)$ be the horizontal subbundle of $TO(M)$. Note that $\mathbf{h}_u(u \cdot): \mathbb{R}^m \xrightarrow{\sim} H_u$ is an isomorphism for each $u \in O(M)$. We give H a Riemannian metric g^H via $g^H = \pi^* g$, i.e., $g_u^H = g_{\pi(u)}(d\pi_u \cdot, d\pi_u \cdot)$. Then

$$dU = \sum_{i=1}^m L_i(U) * dB^i, \quad U_0 = u_0, \quad (6.8)$$

defines horizontal Brownian motion on $O(M)$, and the projection $X = \pi \circ U$ of U down to M is a $BM(M, g)$, started at $x_0 = \pi \circ u_0$. We write $X(x)$ if $\pi \circ u_0 = x$.

If $U_0 = u \in \mathcal{O}(M)$, it has notational advantage to identify B and uB such that the driving Brownian motion B takes its values in $T_{\pi(u)}M = T_xM$ instead of \mathbb{R}^n ; the SDE (6.8) then reads as

$$dU = \sum_{i=1}^m L_i(Uu^{-1}e_i) * dB^i, \quad U_0 = u, \quad (6.9)$$

or equivalently

$$dU = \mathbf{h}_U(//_{0,t} * dB), \quad U_0 = u. \quad (6.10)$$

Using these notations an equivalent to Lemma 2.1 can be formulated as follows.

LEMMA 6.2 *Let $X = \pi \circ U$ be a $\text{BM}(M, g)$ where U is constructed as (strong) solution of (6.8). Let $a(r, \cdot)$ be a transformation as in Lemma 2.1, and suppose that $Y_r = a(r, X_r(x))$ is a local martingale on some stochastic interval. Then*

$$N_r = (da(r, \cdot))_{X_r(x)} (T_x X_r) h_r - a(r, X_r(x)) \int_0^r \langle (T_x X_s) \dot{h}_s, //_{0,s} dB_s \rangle.$$

is a local martingale on this stochastic interval; h denotes an arbitrary bounded adapted process with sample paths $h_\bullet(\omega) \in \mathbb{H}(\mathbb{R}_+, T_x M)$, a.a. ω .

Proof Let $u \in \mathcal{O}(M)$ with $\pi(u) = x$, and think of h taking its values in $H_u \cong T_x M$. By assumption,

$$Y_r = a(r, \pi \circ U_r(u)) = \tilde{a}(r, U_r(u)) \quad (6.11)$$

is a local martingale on some interval $[\sigma, \tau[$. The proof of Lemma 2.1 applies to (6.11) and shows that

$$\begin{aligned} N_r &= (d\tilde{a}(r, \cdot))_{U_r(u)} (T_u U_r) h_r - \tilde{a}(r, U_r(u)) \int_0^r \langle (T_u U_s) \dot{h}_s, h_{U_s}(//_{0,s} dB_s) \rangle_H \\ &= (da(r, \cdot))_{X_r(x)} (T_x X_r) h_r - a(r, X_r(x)) \int_0^r \langle (T_x X_s) \dot{h}_s, //_{0,s} dB_s \rangle_{T_x M}, \end{aligned}$$

is a local martingale on $[\sigma, \tau[$ which gives the claim. \square

7. Gradient estimates for harmonic functions

Cranston [4] used coupling methods to give gradient estimates for harmonic functions. Exploiting directly derivative formulae, like formula (3.1), provides an alternative approach. Based on Lemma 6.1 we may start with the following representation for the gradient of a harmonic function.

THEOREM 7.1 *Let $D \subset M$ be a nonvoid relatively compact open subset with smooth boundary $\partial D \neq \emptyset$, and $\tau(x) = \inf\{t \geq 0 : X_t(x) \notin D\}$ the first exit time of X from D when started at $x \in D$. Let $u \in C(\bar{D})$ be L -harmonic on D . Then*

$$\langle (\text{grad } u)_x, v \rangle = -\mathbb{E} \left[(u \circ X_{\tau(x)}(x)) \int_0^{\tau(x)} \langle W_s(\dot{h}_s), //_{0,s} d\tilde{B}_s \rangle \right] \quad (7.1)$$

for any bounded $\mathcal{F}_\bullet(x)$ -adapted process h such that $h_\bullet \in \mathbb{H}(\mathbb{R}_+, T_x M)$, $h_0 = v$, and $h_s \equiv 0$ for $s \geq \tau(x)$, a.s., with the property that $(\int_0^{\tau(x)} \|\dot{h}_s\|^2 ds)^{1/2} \in L^{1+\varepsilon}$ for some $\varepsilon > 0$.

Using the covariant equation (6.3) it is easy to get norm estimates for $W_r = W_r(v)$. For instance, let $c \geq 0$ be such that the following estimate holds:

$$-\operatorname{Ric}(w, w) + 2 \nabla Z(w, w) \leq c \|w\|^2, \quad w \in T_y M, \quad y \in \bar{D}, \quad (7.2)$$

where $\nabla Z(w, w) = \langle \nabla_w Z, w \rangle$. Then

$$\begin{aligned} \frac{d}{dr} \left\| //_{0,r}^{-1} W_r \right\|_{T_x M}^2 &= 2 \left\langle \frac{d}{dr} //_{0,r}^{-1} W_r, //_{0,r}^{-1} W_r \right\rangle \\ &= 2 \left\langle //_{0,r}^{-1} \left[-\frac{1}{2} \operatorname{Ric}(W_r, \cdot)^\# + \nabla Z(W_r(v)) \right], //_{0,r}^{-1} W_r \right\rangle \\ &= -\operatorname{Ric}(W_r, W_r) + 2 \nabla Z(W_r, W_r). \end{aligned}$$

In other words,

$$\|W_r\|^2 = \|W_0\|^2 + \int_0^r [-\operatorname{Ric}(W_s, W_s) + 2 \nabla Z(W_s, W_s)] ds.$$

Thus, if $\|W_0\| = \|v\| \neq 0$, we get

$$\|W_r\|^2 = \|W_0\|^2 \exp \left(\int_0^r [-\operatorname{Ric}(\hat{W}_s, \hat{W}_s) + 2 \nabla Z(\hat{W}_s, \hat{W}_s)] ds \right) \quad (7.3)$$

where $\hat{W}_r = W_r / \|W_r\|$. Together with (7.2), the last equation gives

$$\|W_r\|^2 \leq \|v\|^2 e^{cr}. \quad (7.4)$$

EXAMPLE 7.2 Let $Z \equiv 0$, and suppose that $\operatorname{Ric} \geq -Cg$ on D for some $C \geq 0$, where g is the Riemannian metric on M . Then, for $r \leq \tau(x)$,

$$\|W_r(v)\| \leq \|v\| e^{1/2 Cr}.$$

Given the situation of Theorem 7.1, we get a straightforward estimate for any nonnegative function $u \in C(\bar{D})$ which is L -harmonic on D as follows:

$$\begin{aligned} |\langle (\operatorname{grad} u)_x, v \rangle|^2 &\leq \mathbb{E} \left[u(X_{\tau(x)}(x))^2 \right] \mathbb{E} \left[\left(\int_0^{\tau(x)} \langle W_s(\dot{h}_s), //_{0,s} d\bar{B}_s \rangle \right)^2 \right] \\ &\leq u(x) \left(\sup_{\partial \bar{D}} |u| \right) \mathbb{E} \left[\int_0^{\tau(x)} \|W_s(\dot{h}_s)\|^2 ds \right] \\ &\leq u(x) \left(\sup_{\partial \bar{D}} |u| \right) \mathbb{E} \left[\int_0^{\tau(x)} \|\dot{h}_s\|^2 e^{cs} ds \right]. \end{aligned}$$

Summarizing this argument, we verified the following general estimate for the gradient of harmonic functions on regular domains in a Riemannian manifolds.

COROLLARY 7.3 Let $u \in C(\bar{D})$ be a nonnegative function which is L -harmonic on D . Let K_Z be the smallest constant such that (7.2) holds. Then

$$|\langle (\operatorname{grad} u)_x, v \rangle| \leq u(x)^{1/2} \left(\sup_{\partial \bar{D}} |u| \right)^{1/2} \left(\inf_h \mathbb{E} \left[\int_0^{\tau(x)} \|\dot{h}_s\|^2 e^{K_Z s} ds \right] \right)^{1/2} \quad (7.5)$$

where the infimum is taken over all bounded $\mathcal{F}_\bullet(x)$ -adapted processes h such that $h_\bullet \in \mathbb{H}(\mathbb{R}_+, T_x M)$, $h_0 = v$, and $h_s \equiv 0$ for $s \geq \tau(x)$, a.s.

We are not going to exploit formula (7.5) here further. For explicit estimates, using the described method, the reader is referred to [15].

8. Concluding remarks

The assumptions of Theorem 2.4 can be slightly weakened when combined with the estimates for the covariant equation (6.3) as given in the previous section. More precisely, we have the following result for heat semigroups associated to (1.1). Now, we assume that M with the induced Riemannian metric is complete.

THEOREM 8.1 *Let $f: M \rightarrow \mathbb{R}$ bounded measurable, $x \in M$, and $v \in T_x M$. Then, for any bounded $\mathcal{F}_\bullet(x)$ -adapted process h with sample paths in $\mathbb{H}(\mathbb{R}_+, T_x M)$ such that $(\int_0^{\tau_D(x) \wedge t} \|\dot{h}(s)\|^2 ds)^{1/2} \in L^1$, and the property that $h(0) = v$, $h(s) = 0$ for all $s \geq \tau_D \wedge t$, the following formula holds:*

$$\langle d(P_t f)_x, v \rangle = -\mathbb{E} \left[f \circ X_t(x) 1_{\{t < \zeta(x)\}} \int_0^{\tau_D(x) \wedge t} \langle (T_x X_s) \dot{h}(s), A(X_s(x)) dB_s \rangle \right]. \quad (8.1)$$

Here $\tau_D(x)$ is the first exit time of $X(x)$ from some open neighbourhood D of x such that $K_Z = \sup\{-\text{Ric}(w, w) + 2 \nabla Z(w, w) : y \in D, w \in T_y M, |w| = 1\}$ is finite, and $\|T_x X_r\| 1_{\{r \leq \tau_D(x)\}} \in L^1$ for each r .

Proof We may assume $f \in bC^1(M)$. Otherwise, we use $P_t f = P_{t-\varepsilon}(P_\varepsilon f)$ to get (8.1) with f replaced by $P_\varepsilon f$ for small $\varepsilon > 0$ and with h replaced by h^ε , see part 2) in the proof of Theorem 2.3. The desired formula then follows as $\varepsilon \rightarrow 0$. Next, by Lemma 6.1,

$$N_r = d(P_{t-r} f)_{X_r(x)} W_r(h_r) - (P_{t-r} f)(X_r(x)) \int_0^r \langle W_s(\dot{h}_s), //_{0,s} d\tilde{B}_s \rangle$$

is a local martingale for $0 \leq r < \tau_D(x) \wedge t$. Since $\|W_r(v)\| \leq \|v\| e^{K_Z r/2}$ for $r \leq \tau_D(x)$, we conclude that $(N_{r \wedge \tau_D(x)})$, $r \in [0, t]$, is already a martingale under the given assumptions; on a complete Riemannian manifold $d(P_s f)_x$ is bounded for $s \leq t$, $x \in M$, e.g. [2] or [15]. This implies

$$\langle d(P_t f)_x, v \rangle = -\mathbb{E} \left[f(X_t(x)) 1_{\{t < \zeta(x)\}} \int_0^{\tau_D(x) \wedge t} \langle W_s(\dot{h}(s)), //_{0,s} d\tilde{B}_s \rangle \right]. \quad (8.2)$$

Note that, since $A(X_s(x)) dB_s = //_{0,s} d\tilde{B}_s$ and $W_r(v) = \mathbb{E}^{\mathcal{F}_r(x)}[(T_x X_r) v 1_{\{r \leq \tau\}}]$, we are able to recover (8.1) from (8.2). \square

We remark that if K_Z is finite for $D = M$, and $\zeta(x) = \infty$, a.s., then (8.2) holds with $\tau_D(x) \equiv \infty$. Note that, if the drift Z is $\text{grad } \varphi$ for a smooth function φ , then $K_Z < \infty$ implies infinite lifetime, see [1].

Acknowledgements. The author wishes to express his thanks to K. D. Elworthy, Xue-Mei Li and J. Zabczyk for encouraging and helpful discussions on the subject. Finally, he would like to thank David Elworthy and the University of Warwick for the warm hospitality during his stay at Warwick where this work was stimulated.

References

- [1] D. BAKRY. Un critère de non-explosion pour certaines diffusions sur une variété riemannienne complète. *C. R. Acad. Sci. Paris, Série I* **303** (1986) 23–26.

- [2] D. BAKRY and M. LEDOUX. Levy-Gromov's isoperimetric inequality for an infinite dimensional diffusion generator. *Invent. math.* **123** (1996) 259–281.
- [3] J. M. BISMUT. *Large deviations and the Malliavin calculus*. Birkhäuser: Progress in Mathematics **45** (1984).
- [4] M. CRANSTON. Gradient estimates on manifolds using coupling. *J. Funct. Anal.* **99** (1991) 110–124.
- [5] K. D. ELWORTHY. Geometric aspects of diffusions on manifolds. In: P. L. Hennequin (Ed.) *École d'Été de Probabilités de Saint-Flour XV–XVII*. Lect. Notes in Math. **1362** (1988) 277–425.
- [6] K. D. ELWORTHY and X.-M. LI. Formulae for the derivatives of heat semigroups. *J. Funct. Anal.* **125** (1994) 252–286.
- [7] K. D. ELWORTHY and X.-M. LI. A class of integration by parts formulae in stochastic analysis. In: N. Ikeda et al. (Eds.) *Itô's Stochastic Calculus and Probability Theory*. New York: Springer-Verlag, 1996, 15–30.
- [8] K. D. ELWORTHY, X.-M. LI and Y. LE JAN. Concerning the geometry of stochastic differential equations and stochastic flows. In: K. D. Elworthy, S. Kusuoka and I. Shigekawa (Eds.) *New Trends in Stochastic Analysis*. Proc. Taniguchi Symposium 1994. World Scientific Press, 1997.
- [9] K. D. ELWORTHY and M. YOR. Conditional expectations for derivatives of certain stochastic flows. In: J. Azéma et al. (Eds.) *Séminaire de Probabilités XXVII*. Lect. Notes in Math. **1557** (1993) 159–172.
- [10] M. EMERY. *Stochastic Calculus in Manifolds* (with an appendix by P. A. Meyer). Berlin: Springer, 1989.
- [11] N. V. KRYLOV. Quasiderivatives for solutions of Itô's stochastic equations and their applications. In: T. Lindstrøm et al (Eds.) *Stochastic Analysis and Related Topics*. Proc. of the Fourth Oslo-Silivri Workshop on Stochastic Analysis, Oslo 1992. Stochastic Monographs **8**. Gordon and Breach Sci. Publishers, 1993, 1–44.
- [12] H. KUNITA. *Stochastic flows and stochastic differential equations*. Cambridge: Cambridge University Press, 1990.
- [13] X.-M. LI. Stochastic differential equations on noncompact manifolds: moment stability and its topological consequences. *Probab. Theory Relat. Fields* **100** (1994) 417–428.
- [14] X.-M. LI. Strong p -completeness of stochastic differential equations and the existence of smooth flows on noncompact manifolds. *Probab. Theory Relat. Fields* **100** (1994) 485–511.
- [15] A. THALMAIER and F.-Y. WANG. Gradient estimates for harmonic functions on regular domains in Riemannian manifolds. Warwick Preprints: 32/1996.

Anton Thalmaier
 NWF I – Mathematik
 Universität Regensburg
 D – 93040 Regensburg
 GERMANY

E-mail: anton.thalmaier@mathematik.uni-regensburg.de